

CLAIMS

What is claimed is:

1. A method for markup language document compression, comprising:
defining a schema that specifies the structure of a markup-language conforming document;
defining in said schema the types of elements and attributes that comprise said conforming document;
assigning names in said schema for each of said elements and attributes of said document;
defining relationships between said elements;
defining relationships between said attributes and said elements;
assigning a token in said schema representing each said element name and each said attribute name of said document; and
replacing in said document each said element name and each said attribute name with said assigned token.
2. The method of claim 1 further comprising declaring a namespace in said schema for said tokens.
3. The method of claim 1 wherein numbers and alphabetic characters are used for tokens.
4. The method of claim 3 wherein said numbers are used for said tokens representing said element names and said alphabetic characters are used for said tokens representing said attribute names.
5. The method of claim 1 wherein said assigning comprises parsing said schema.
6. The method of claim 1 further comprising storing said document in a computer-readable storage medium.
7. The method of claim 1 further comprising storing said schema in a computer-readable storage medium.

8. The method of claim 7 further comprising storing said document in a computer-readable storage medium separately from said schema.

9. The method of claim 7 further comprising transmitting said document separately from said schema.

10. The method of claim 9 further comprising:
receiving said separately transmitted document by an application; and
retrieving token information in said schema by accessing said separately stored schema.

11. The method of claim 10 further comprising accessing said stored schema by said application from a location selected from the group consisting of known locations and locations defined in said document.

12. The method of claim 11 further comprising replacing each token with an assigned element and/or attribute name.

13. The method of claim 12 comprising parsing said schema.

14. The method of claim 1 wherein said markup language comprises eXtensible Markup Language (XML).

15. A system for markup language document compression, comprising:
means for defining a schema that specifies the structure of a markup-language conforming document;
means for assigning names in said schema for each of said elements and attributes of said document;
means for assigning a token in said schema representing each said element name and each said attribute name of said document; and
means for replacing in said document each said element name and each said attribute name with said assigned token.

16. The system of claim 15 further comprising means for parsing said markup-language conforming document.

17. The system of claim 15 wherein numbers and alphabetic characters are used for tokens.

18. The system of claim 15 further comprising means for storing said document in a computer-readable storage medium.

19. The system of claim 18 further comprising means for storing said schema separately from said document in a computer-readable storage medium.

20. The system of claim 19 further comprising means for transmitting said document separately from said schema.

21. The system of claim 20 further comprising:
means for receiving said separately transmitted document by an application; and
means for replacing said tokens with said assigned element and/or attribute names.

22. The system of claim 15 wherein said markup language comprises eXtensible Markup Language (XML).

23. A processing system comprising:
a markup-language conforming document, said document having elements and attributes such that each element name and each attribute name is represented in said document by a token, thereby compressing said document;
a schema that specifies the structure of said document, said schema defining said token representing each said element name and each said attribute name;
a software application operable to access said schema; and
a processing element communicatively coupled with said software application, said processing element operable to retrieve said document and operable to process said document in conjunction with said software application.

24. The processing system of claim 23 further comprising a memory unit communicatively coupled with said software application and said processing element, said memory unit operable to provide said document to said software application and said processing element.

25. The processing system of claim 23 further comprising a storage facility operable to provide said schema to said software application.

26. The processing system of claim 23 further comprising a document destination facility incorporating a document translator.

27. The processing system of claim 26 wherein said document translator is operatively coupled with a document parser.

28. The processing system of claim 23 wherein said markup language comprises eXtensible Markup Language (XML).

29. Computer-executable software code stored to a computer-readable medium, said computer-executable software code comprising:

code for using a schema-defined markup language conforming document, wherein element names and attribute names of said document are replaced with tokens assigned and recorded in said schema.

30. The computer-executable software code of claim 29, further comprising code for storing said document to a computer-readable storage medium.

31. The computer-executable software code of claim 29, further comprising code for storing said schema to a computer-readable storage medium separately from said document.

32. The computer-executable software code of claim 31, further comprising code for transmitting said document separately from said schema.

33. The computer-executable software code of claim 32, further comprising:
code for receiving said separately transmitted document by an application; and
code for replacing said tokens with said assigned element and/or attribute names.

34. The computer-executable software code of claim 29 wherein said markup
language comprises eXtensible Markup Language (XML).